Estimating Policy Effects with Interrupted Time Series Following a Nuisance Interruption

NESS 2025

Benjamin Stockton Linda Kahn Shilpi Mehta-Lee
Erinn Hade

Department of Population Health, New York University Grossman School of Medicine

2025-06-02

Introduction

The Problem

- Public policy analysis is challenging.
  - Lack of comparison groups and no randomization
  - Limited sample sizes
  - A strong desire for quick evaluations
- Compounded by the presence of confounding events.
  - Pandemics or other public health events, e.g. COVID-19
  - Changes in federal or state policy simultaneously
  - Changes in NGO policy and guidelines, e.g. AHA, ACOG, etc.

Our Question

What happened to women's reproductive health in Texas after their six-week abortion ban?

Our Innovation

Innovation

Address **nuisance** or **confounding interruptions** in interrupted time series analysis (ITS) by modeling the nuisance with a range of predictors across several models and using Bayesian Model Averaging (BMA) to infer policy effects on a particular outcome.

Limitations

- Confounding event occurs prior to the policy implementation.
- Single-unit time series without a control.

Data

- Electronic medical records are aggregated over two month periods from Epic Cosmos from January 1, 2017 through June 30, 2023.
- Diagnostic code frequencies were reported for pregnant women along with the total number of pregnant women and the total number of women between ages 10 and 55.
- Frequencies are reported for pregnancies that began in that two month period.
- Outcomes: Miscarriage, Hypertensive Disorders of Pregnancy, Depression, Fertility, etc.

Interrupted Time Series

- ITS analysis has been used to assess policy impacts in public health and policy research.[a]
- Prior work has used segmented regression, Bayesian structural time series (BSTS), and ARIMA models.
- Quantities of interest: point-wise differences and cumulative differences.

[a]See Lopez Bernal et al. [6]; Gianacas et al. [5]; Papadogeorgou et al. [7].

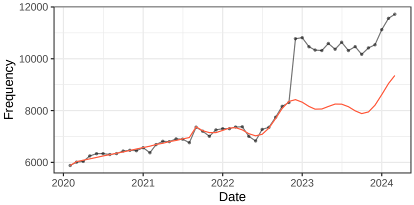


Figure 1: An example of a simulated interrupted time series with $T = 50$.

Methods

Causal Assumptions

Standard causal assumptions:

1. Consistency - Observed is a faithful observation of the potential outcome.
2. Interference - There's a single unit, so treatment on another unit cannot affect it.
3. No unmeasured confounding - No confounding events start simultaneously with the policy.

Building on previous work[1]:

1. Potential outcomes pre-intervention are unaffected by the intervention.
2. Covariates are independent of the intervention and the outcome is stationary conditional on the covariates.

[1] Papadogeorgou et al. [7]

Poisson Interrupted Time Series Model

- Let the outcome be $Y_t | \beta, \phi \sim Pois(\mu_t)$.
- Set priors on $\beta_k \overset{iid}{\sim} N(0, 10^2)$ and $\phi_p \overset{iid}{\sim} N_{[-1,1]}(0, 1)$.
- $f(t, t_C, C_t; \beta)$ is a function of the confounder.
- A_t is an indicator for the start of the intervention and t_A is the time the intervention occurred.
- $\mathbf{x}_t$ is a vector of exogenous predictors. W_t is an offset.

Fully Parametric Model

$$\log(\mu_t) = \beta_0 + \mathbf{x}_t'\beta_1 + f_C(t, t_C, C_t; \beta) + \beta_3 A_t + \beta_4 (t - t_A) A_t + \sum_{p=1}^{P} \phi_p \log(Y_{t-p} + 1) + \log(W_t), \quad (1)$$

Fit this model on all of the data pre- and post-intervention.

Semi-Parametric Model

The semi-parametric model drops the term $\beta_3 A_t + \beta_4(t - t_A)A_t$,

$$\log(\mu_t) = \beta_0 + \mathbf{x}_t'\beta_1 + f_C(t, t_C, C_t; \beta) + \sum_{p=1}^{P} \phi_p \log(Y_{t-p} + 1) + \log(W_t), \tag{2}$$

Fit only on the pre-intervention data in order to make a counterfactual forecast.

- Fit using Stan[2] and CmdStanr.[3]

[2] Team [8]

[3] Gabry et al. [4]

Parametric vs Semi Parametric Approach

Table 1: Comparison of approaches.

Fully Parametric	Semi-Parametric
In-sample prediction	Out-of-sample forecast
Parametric form for policy effect	No specified form for policy effect
Confounder effect is averaged over selected parametric forms	Confounder effect is averaged over selected parametric forms

Confounding Nuisances

- Let $C_t = I(t \geq t_C)$ be an indicator for the confounding period.
- The effect the confounder on the outcome may be uncertain, so several impact shapes could be considered.

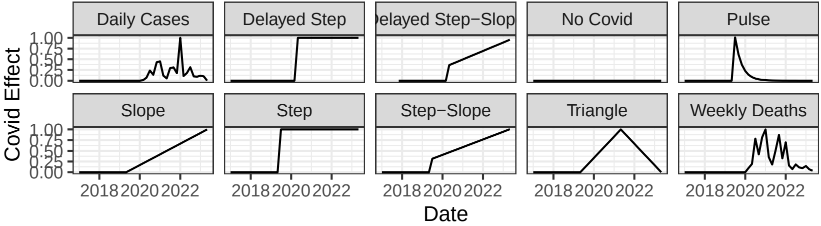


Figure 2: Potential shapes for the Covid-19 impact.

Bayesian Model Averaging

1. Fit each model with selected confounder forms.
2. Score each model's out-of-sample predictive performance on the pre-intervention data using Leave-Future-Out CV on expected predictive log-likelihood.[4]
3. Use scores to draw weighted samples from the MCMC posterior predictive draws for each model.
4. Perform Bayesian inference with the weighted sample of posterior predictive draws as usual.

[4] Bürkner et al. [3] and Barigou et al. [1].

Simulation

Set-Up

Quantity of Interest: Cumulative Difference between Observed and True Counterfactual.

Table 2: Simulation Settings

Property	Values
Series Length: T	39, 50
Interruption Timing: t_A	Mid ($T/2$), Late ($3T/4$)
Confounding Shape: C	Simple (step), Complex (sine wave)
Length of Confounding	Short ($t_A - t_C = 5$), Long ($t_A - t_C = 10$)
Strength of Confounding: β_2	None 0, Weak 0.01, Strong 0.1
Strength of Intervention: α	None 0, Strong 0.1

Models

Table 3: Model averaging over no confounding, step, step-slope, and sine predictor models.

Name	Approach	Abbreviation
Bayesian Model Averaging	Semi and Fully Parametric	(SP BMA, FP BMA)
CausalImpact[5]	Bayesian Structural Time Series	(BSTS)
No Confounding	Semi and Fully Parametric	(SP No Conf, FP No Conf)
Step Slope Confounding	Semi and Fully Parametric	(SP Step-Slope, FP Step-Slope)

[5] Brodersen et al. [2]

Results

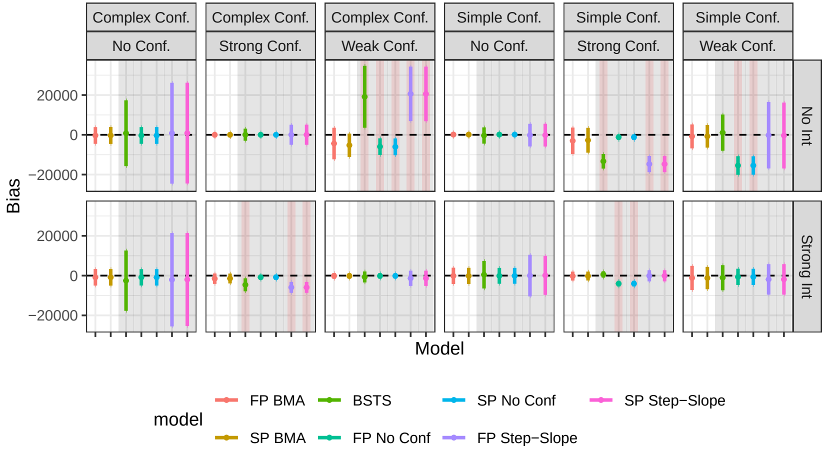


Figure 3: Bias in Cumulative Difference for select simulations.

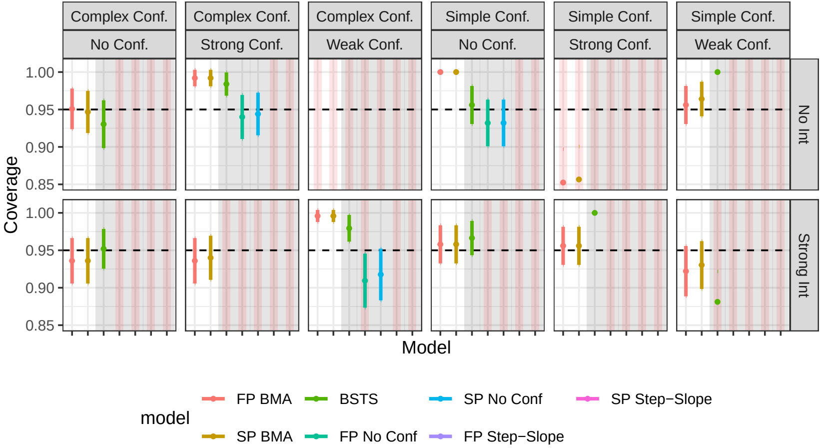


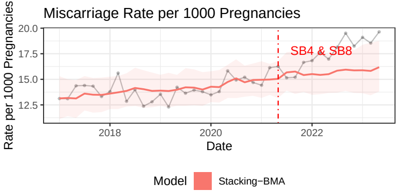
Figure 4: Credible interval coverage of Cumulative Difference for select simulations.

Case Study: Texas Abortion Ban

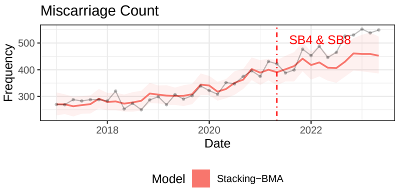
Policy Background & Context

- The Texas legislature passed a six-week abortion ban through SB4 and SB8 in July 2021.
- The ban was implemented in September 2021, so pregnancies that began in July 2021 or later would be potentially impacted by the ban.
- The Dobbs v Jackson decision (June 2022) by the US Supreme Court overturned the Roe v Wade precedent that protected the right to abortion.

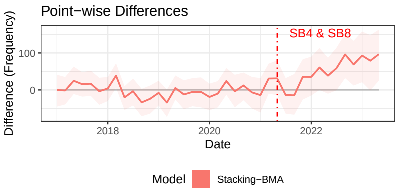
Results - Miscarriage



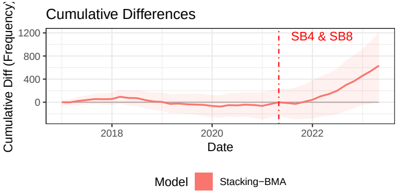
(a)



(b)

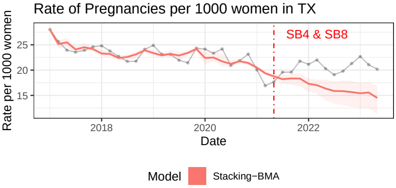


(c)

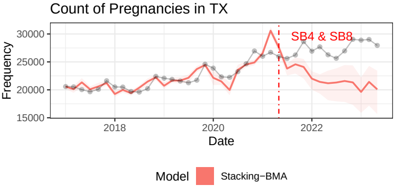


(d)

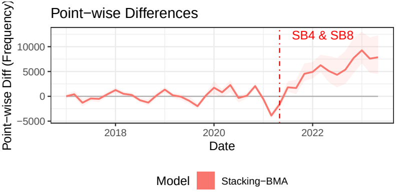
Results - Fertility



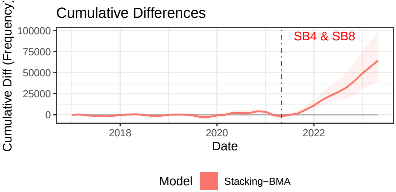
(a)



(b)



(c)



(d)

Discussion

Conclusion

- In our analysis, we found that miscarriages and the number of pregnancies in Texas both increased following the abortion ban in 2021.
- Our analysis is restricted by the use of Epic Cosmos data which is not representative of all Texas women.
  - How can we generalize our model from this setting to the full population?
- Alternatives to single unit ITS:
  - Synthetic Controls, Difference in Difference, Bayesian structural time series (BSTS)

Acknowledgements

- Research was supported in part by The Educational Foundation of America and by NYU CTSA grant UL1 TR001445 from the National Center for Advancing Translational Sciences, National Institutes of Health.
- The computational work performed on this project was done with help from the NYU Big Purple High-Performance Computing cluster.

Questions?

References I

[1] Karim Barigou, Pierre-Olivier Goffard, Stéphane Loisel, and Yahia Salhi. Bayesian model averaging for mortality forecasting using leave-future-out validation. *International Journal of Forecasting*, 39(2):674–690, April 2023. ISSN 0169-2070. doi: 10.1016/j.ijforecast.2022.01.011.

[2] Kay H. Brodersen, Fabian Gallusser, Jim Koehler, Nicolas Remy, and Steven L. Scott. Inferring causal impact using Bayesian structural time-series models. *The Annals of Applied Statistics*, 9(1):247–274, March 2015. ISSN 1932-6157, 1941-7330. doi: 10.1214/14-AOAS788.

[3] Paul-Christian Bürkner, Jonah Gabry, and Aki Vehtari. Approximate leave-future-out cross-validation for Bayesian time series models. *Journal of Statistical Computation and Simulation*, 90(14):2499–2523, September 2020. ISSN 0094-9655. doi: 10.1080/00949655.2020.1783262.

[4] Jonah Gabry, Rok Češnovar, and Andrew Johnson. *cmdstanr: R Inferfacet to 'CmdStan'*. 2023. URL https://mc-stan.org/cmdstanr/,https://discourse.mc-stan.org.

References II

[5] Christopher Gianacas, Bette Liu, Martyn Kirk, Gian Luca Di Tanna, Josephine Belcher, Suzanne Blogg, and David J. Muscatello. Bayesian structural time series, an alternative to interrupted time series in the right circumstances. *Journal of Clinical Epidemiology*, 163:102–110, November 2023. ISSN 0895-4356. doi: 10.1016/j.jclinepi.2023.10.003.

[6] James Lopez Bernal, Steven Cummins, and Antonio Gasparrini. Interrupted time series regression for the evaluation of public health interventions: A tutorial. *International Journal of Epidemiology*, 46(1):348–355, 2017. ISSN 0300-5771, 1464-3685. doi: 10.1093/ije/dyw098.

[7] Georgia Papadogeorgou, Fiammetta Menchetti, Christine Choirat, Jason H. Wasfy, Corwin M. Zigler, and Fabrizia Mealli. Evaluating federal policies using bayesian time series models: estimating the causal impact of the hospital readmissions reduction program. *Health Services and Outcomes Research Methodology*, 01 2023. doi: 10.1007/s10742-022-00294-8. URL https://doi.org/10.1007/s10742-022-00294-8.

References III

[8] Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual*. 2023. URL https://mc-stan.org.